The Ontario Secondary School Literacy Test

Resource Guide
2000–2001

EQAO
Education
Quality and
Accountability
Office

Table of Contents

1. Introduction to the Education Quality and Accountability Office (EQAO) 1
2. Purpose of the Resource Guide 1
3. Uses of Test Results 1
4. The Individual Student Report 2
5. Test Background 2
6. Test Development and Administration 3
7. The Reading Component 3
8. The Writing Component 4
9. Marking and Reporting the Reading Component 5
10. Marking and Reporting the Writing Component 5
11. Setting the Standards in Reading and Writing for Pass/Fail 7
12. Student Background Information 9

Appendix A
An Overview of the Foundations for the Ontario Secondary School Literacy Test 12

Appendix B
Grade 10 Test of Reading and Writing — A Report on the Process by Robert J. Wilson 19

Appendix C
Validation of Standards, August 2000 29

Appendix D
Procedures for the Implementation of Accommodations for Students with Special Needs 31

Appendix E
Special Provisions for Students Enrolled in ESL/ELD Programs 40

Appendix F
Recommended Resources for the Ontario Secondary School Literacy Test 42

1. Introduction to the Education Quality and Accountability Office (EQAO)

EQAO promotes greater accountability and improved student achievement in Ontario's publicly funded school system. An independent, arm's-length agency of the provincial government, EQAO provides parents, teachers and the public with accurate and reliable information on which to base decisions about improvement of student achievement.

EQAO values

- the well-being of learners above all other interests.
- only that information that has the potential to bring about constructive change and improvement.
- the dedication and expertise of Ontario's educators and their involvement in all of its activities.
- the delivery of its programs and services with equivalent quality in English and French

EQAO's roles and responsibilities are to

- measure students' achievement against the expectations in the provincial curriculum.
- conduct province-wide assessments of elementary and secondary school student achievement in reading, writing and mathematics in Grades 3 and 6, mathematics in Grade 9, and reading and writing in Grade 10.
- coordinate Ontario's participation in national and international assessments.
- analyze data on student achievement and the learning environment to determine where progress is being made and where gaps still need to be closed.
- develop and implement a provincial Education Quality Indicators Program (EQUIP).
- make recommendations for school and system improvement.
- provide opportunities to help teachers enhance their skills and expertise in assessment.
- report to parents, educators and the public about student achievement and education quality.
- monitor improvements in student achievement over time to reinforce best practices and to identify areas where further attention is required.
- support a culture of assessment through ongoing testing, research and implementation of a quality indicators program.

2. Purpose of the Resource Guide

This guide is intended to provide information on the Ontario Secondary School Literacy Test. It provides the background necessary to understand all aspects of the test and will assist students, parents, educators and the public in interpreting the individual, school and board results.

3. Uses of Test Results

The primary focus of the test design, scoring and reporting is on the individual student. Students whose work has met the standard established for this test are reassured that they have demonstrated reading and writing skills that are sufficiently developed for the end of Grade 9. With appropriate further effort and development they should be successful in acquiring the reading and writing skills required to complete secondary school programs. For students who do not successfully complete this

test, the test achievement data as reported on the *Individual Student Report* deliver diagnostic information for remediation.

The test was developed from the expectations for reading and writing across the curriculum. Therefore, individual schools may find the achievement data helpful in reviewing the implementation of cross-curricular reading and writing expectations in the school. As well, individual school boards may find the achievement data give an indication of program needs, which will provide direction on ways to focus cross-curricular literacy initiatives and to further the implementation of the curriculum expectations in all subjects related to literacy.

The test was not designed to produce data on which to base comparisons of school performances. There is not sufficient contextual information to qualify and make equitable inter-school comparisons of achievement data. Therefore, it is EQAO's position that the use of test results to rank schools is misleading and inappropriate. Rankings lead people to draw sweeping conclusions or make simplistic comparisons that fail to consider all of the factors that affect achievement. In the absence of contextual data, which would minimize the possibility of misinterpretation of the results, the comparison of schools may lead to erroneous conclusions and is, therefore, counterproductive and unfair.

4. The Individual Student Report

- The *Individual Student Report* is intended to provide feedback for students and to engage parents in supporting improvement, particularly for students who were unsuccessful on this test administration.
- Each student who participated in the October 2000 Grade 10 test will receive an *Individual Student Report*. The results of the October 2000 test will not be recorded on the students' Ontario Student Transcripts or on their report cards.
- Students who successfully completed the test demonstrated different degrees of competence ranging from acceptable to outstanding. All students who were successful should continue to work with their teachers to improve these skills as they progress through high school.
- Only unsuccessful students will receive feedback on their test performance. Unsuccessful students are urged to use the feedback from this test to discuss with their teachers how to improve their skills.
- To protect the confidentiality of students and prevent any potential marker bias, student booklets are identified only by a bar-coded identification system. For reporting purposes, student results are matched to a student name based on information provided by schools. In some cases the information necessary to carry out this matching process was either missing or ambiguous. EQAO is attempting to resolve such issues with school administrators. Where this process is not complete at the time that reports are being produced, the student report will include the notation "data issues yet to be resolved, *ISR* may be issued at a later date" and no results will be shown.

5. Test Background

- In January 1998, the Ministry of Education announced that Ontario students would have to pass the Ontario Secondary School Literacy Test as a graduation requirement. The test would be administered in Grade 10 in order to provide opportunities for students who did not pass to improve their skills and retake the test in later grades.
- This diploma requirement will come into effect in the 2001–2002 academic year for students who enrolled in Grade 9 in September 2000 and thereafter. The successful

completion of the Grade 10 test is not a graduation requirement for students who entered Grade 9 in September 1999.

- Those students who were unsuccessful on the October 2000 administration cannot retake the test. Similarly, students who were exempted, deferred or absent for the October 2000 test cannot take the test in the future.
- The test measures students' reading and writing skills based on language expectations common to all subjects in the provincial curriculum up to the end of Grade 9. Students must pass both the reading and writing components to successfully complete the test.
- For information on the test framework and the assumptions on which the test is based, please see Appendix A.

6. Test Development and Administration

- Over 300 Ontario teachers and assessment specialists participated in the various stages of test development.
- Developers designed a blueprint for what the test would assess and then selected reading passages, and constructed reading questions and writing tasks according to the blueprint specifications. This was to ensure that all future tests will be comparable over time.
- English-language and French-language tests were developed in parallel and with frequent consultation, but by separate teams of developers.
- Questions were piloted twice with Ontario students. Pre-pilot and pilot testing were undertaken to determine whether questions

  > were at an appropriate level of difficulty;

  > were age- and grade-appropriate;

  > measured a sufficient range of skills; and

  > contained clearly stated instructions.
- Based on the results of the pilots, some questions were eliminated and others were revised for inclusion in the field test.
- The revised questions were compiled for field testing with a demographically representative sample of students across the province and the test was again revised based on student performance data.
- The pilots and field tests were also used to establish appropriate test administration conditions that would ensure consistent implementation across the province.
- The test was composed of a reading and writing component. The reading questions and writing tasks were administered over a total of five hours (including breaks) on two consecutive half days, October 11 and 12, 2000.
- Sample reading questions, writing tasks and samples of successful student work from the October 2000 test will be posted on the EQAO Web site at www.eqao.com in March 2001.

7. The Reading Component

In the reading component, students were required to read 12 reading selections and answer 100 questions about what they read.

There were three types of reading selections:

- Information (e.g., explanation, instructions)
- Graphic (e.g., schedules, graphs, tables)
- Literary (e.g., story, description, dialogue)

The reading selections represented a range of difficulty. The factors determining the range of difficulty included the following:

- students' familiarity with the subject matter
- complexity of language, structure and thought
- length

The test used three types of questions:

- multiple choice
- short answer (requiring students to write their own responses in a word or a few words)
- short answer with explanation (requiring students to justify their answers or explain their thinking in a sentence or two)

Test questions were designed at three levels of difficulty, and validated by student responses during the pilots and field tests. The factors determining the range of difficulty included the following:

- the type, quality and quantity of information required
- the scope of thinking required
- the complexity of the question

Reading Skills

The reading component measured the students' demonstration of three reading skills essential to understanding:

- **understanding** ideas and information stated directly in the reading selection
- **understanding** indirectly stated ideas and information (making inferences about what the selection means)
- **making connections** between personal knowledge and experiences and the ideas and information in the reading selections (interpreting meaning)

Reading Strategies

In addition, results in four supporting reading strategies will be reported at the school, board and provincial levels.

- **vocabulary**: using knowledge or context to understand familiar, unfamiliar and technical vocabulary and imaginative uses of language (such as idiomatic expressions, colloquial language, imagery, metaphoric expressions, etc.)
- **syntax**: using grammatical and syntactic structures (such as subject, predicate, sentence types, verb tenses, punctuation, etc.)
- **organization**: using knowledge of overall structural and organizational elements (such as paragraph structure, topic sentence, point of view, cause and effect, etc.)
- **graphic features**: using knowledge of visual features (such as titles, illustrations, diagrams, charts, captions, layout, etc.)

8. The Writing Component

In the writing component, students were asked to write four pieces:

- a summary
- a series of paragraphs expressing an opinion
- a news report
- an information paragraph

Writing Skills

The writing component measured students' demonstration of the following six skills across the four pieces of writing:

- developing a main idea
- providing supporting details
- organizing and linking ideas and information

- using appropriate tone for the purpose and intended readers
- using correct grammar and punctuation
- using correct spelling

It is important to take into consideration the difference between process writing during classroom instruction and on-demand writing under test conditions. In most classroom writing assignments, students follow a process that includes access to dictionaries, peer and teacher feedback and opportunities to revise and edit over time. This test requires students to write without these aids.

9. Marking and Reporting the Reading Component

- Students' written responses were assessed against an answer key and marked as either correct or incorrect. The short-answer-with-explanation questions could also be marked as partially correct.
- The development of this answer key required the analysis of hundreds of student test booklets that demonstrated a range of answers. From that range, examples of correct, incorrect or partially correct responses were selected for inclusion in the answer key. Markers matched individual student responses to the answer key to determine which score to assign.
- To ensure the highest possible degree of fairness and consistency in marking, the four different portions of the reading component were each marked by a different person.
- In cases where a student's total mark was close to the pass/fail line, a second set of four markers re-marked the material and the student's final score was based on the average of the two sets of marks.
- About 25% of student work went through this second phase of marking. The final mark for these students was based on the assessment of eight markers, combined with the computer-generated multiple-choice score.

Calculating the Reading Score

- The reading score for students, which determined whether or not they passed this component, depended on the total number of questions answered correctly.
- In this calculation, multiple-choice and short-answer questions contributed equally to the overall reading score and questions deemed partially correct were awarded half the value of a correct answer.
- Questions left blank and illegible answers were treated the same as incorrect answers.

Student Feedback on Reading

- Although the standard for passing or failing the reading component is based on a total score, for students who did not pass, feedback is provided on the individual student's relative strength and weakness on the three reading skills and the three types of reading selections.
- Only students who were not successful receive detailed feedback.

10. Marking and Reporting the Writing Component

- Scoring the writing component followed a two-phase process. In the first phase, each of the writing tasks was marked holistically. That is, markers assigned a holistic mark taking into account all aspects of the passage: how well it addressed the assigned task, clarity of communication, organization, grammar and punctuation, and spelling. Each marker was trained to mark one task

only, so each student's entire work was marked by four different markers during phase-one marking.

- To ensure consistency in scoring, markers used a scoring rubric unique to this test. The holistic writing rubric described student work on each column of the rubric and was supported by samples of student work that exemplified the data for each of the four columns.
- The specific criteria for assigning marks varied from task to task. No marks were given for tasks not attempted or illegible. Generally the following descriptions applied:
  - > one mark for work that provided insufficient information (off topic or inconsistent with the requirements of the task) or for work with grammatical or spelling errors that prevented clear communication.
  - > two marks for work that was on task but still had significant shortcomings, such as insufficient information, errors that interfered with clear communication, or faulty structure
  - > three marks for work that successfully addressed the assigned task with minor shortcomings in such areas as organization, grammar and punctuation, or spelling
  - > four marks for well-developed and effective student writing with few, if any, grammatical or spelling errors
- The second phase of marking for the writing component of the test was analytic marking. As with the holistic rubrics, the analytic rubrics were unique to this test and were illustrated by samples of student work.
- In the analytic scoring phase, each task was scored on six characteristics. Two markers marked each task analytically. The first analytic marker marked main idea, supporting details and organization. The second analytic marker marked grammar and punctuation, spelling and tone. This phase provided additional information for making the final determination of whether students were successful or not on the writing component and also provided the basis for feedback for students who were not successful.

Calculating the Writing Score

- The overall writing score was determined by combining all the holistic scores and all the analytic scores through a formula derived as part of the standard-setting process. In this formula, the weighting of the holistic scores is roughly equal to the weighting of the combined analytic scores.
- Not every student's work went through both marking phases. If, after phase one, it was clear from the holistic marks that the work was of sufficient quality to meet the standard, phase-two marking was not required. For these cases, the analytic marks were not required to confirm success or to provide specific feedback.
- Similarly, students who attempted only one or two of the tasks did not have their work marked analytically. In these cases, there was not enough work to give meaningful and reliable feedback on the six writing characteristics.
- Over 50% of the student writing was marked both holistically and analytically. For these students, the report is based on the scores of 12 trained markers.
- Many students lost significant potential marks either because they did not complete some of the tasks or because the work they provided was not consistent with the form or topic of the given task. No marks were awarded for tasks that were not attempted or were illegible. Work that did not address the assigned task received only one mark in

the holistic phase of the marking and in some cases could not be marked analytically.

- Examples of such "non-compliant" writing include the following:
  - > submitting a single paragraph in response to a task that requested "a series of paragraphs (a minimum of three)..."
  - > submitting a passage on the topic of "Dangerous Driving" in response to an assigned topic of "Teenage Fitness"
  - > submitting a personal opinion on a topic when the assigned task was to write a summary of given information
- The importance of following the instructions was emphasized in the *Support Materials*, in the test orientation sessions and in the actual test instructions. Students were advised in advance that if their written work did not correspond to the instructions, their work would not be fully scored.

Student Feedback on Writing

Students who were not successful on the writing component of the test are given two kinds of feedback in two tables on the *Individual Student Report*.

- The first table describes student work on each of the four tasks. This section does not address the quality of the writing. Rather, this section addresses whether the task was attempted, whether the work was consistent with the tasks as required and why, in some cases, there was not sufficient information to report. This feedback tells students whether they attempted each task, whether their work was legible, whether their work was on topic and whether they addressed the task they were required to do.
- The second table describes the student's written work in terms of the six characteristics: main idea, supporting details, organization, tone, grammar and punctuation, and spelling.
- If students who were unsuccessful wrote on task and on topic, their work was scored during phase-two analytic marking in terms of those six characteristics.
- Unless the student completed at least three tasks **as assigned**, reliable feedback could not be provided for this section.
- Students with two or more tasks left blank or whose work was deemed non-compliant would receive the comment "Insufficient work to mark...".
- Only students who did not pass the writing component received detailed feedback.

11. Setting the Standards in Reading and Writing for Pass/Fail

The standards were established by parallel English- and French-language committees of both educators and members of the public. Standard-setting involved a three-stage process:

1. Markers assigned numerical scores to student work from the field test.
2. Without reference to those scores, standard-setting committees determined which student work was acceptable and which was not, in order to develop the standards for acceptable work.
3. Assessment experts correlated the numerical scores and the standards to calculate the cut score for pass/fail.

Participants

In the English-language and French-language committees, there was equal representation of educators currently employed in the elementary/secondary school system and members from the public at large.

Directors of Education nominated educators. The members of the public were nominated by senior education officials and community

leaders on the basis of their experience with schools and students, or their familiarity with environments in which reading and writing skills are required. Participants were not designated as representatives of any groups or organizations to which they might belong.

The diversity of the province was represented, insofar as that was possible in 18-member committees.

The Process

The nine English-language educators and nine members from the public came together on May 30, 2000, and worked for three days. There was a parallel French-language committee that met June 13 to 15, 2000.

Educators familiar with large-scale assessment worked with the committees until the final stage of decision-making, at which point they turned the process over to the committees, which then functioned independently. Staff and consultants from EQAO supported and recorded the process.

For the English-language sessions, assessment experts from three universities (University of British Columbia, Queen's University and the Université d'Ottawa) and experts with experience in standard-setting for national and international assessments provided advice to ensure the formal validity and reliability of the standards. Comparable assessment expertise was in place for the French-language sessions.

For each session, the 18 committee members were presented with background information, materials relevant to the Grade 10 test and demographic information about Ontario students.

The committee reviewed a wide range of student work from the March 2000 Grade 10 field test. The papers had already been marked and each question had been assigned a numerical score (although neither these scores nor the distribution of the scores across the population were known to the committee members). The committees developed descriptions of students' performance for reading and writing at the unacceptable and acceptable (incomplete and complete) levels of proficiency. The committees then applied these descriptors in reviewing student test performance in reading and writing to determine which student papers should pass and which should not.

The descriptions of student work below were created by the English-language committee and the wording has not been changed except to correct small mechanical errors. Many of the words and phrases were subject to lengthy discussion and occasionally to votes of the members.

For purposes of sorting student work, the committee also developed a description of work that was "advanced," well beyond the standard required for a pass. Since the test will be reported only as pass/fail, the third category is not used in the reporting process.

Standard for Reading

Incomplete

The student reads with limited accuracy and proficiency:

- demonstrates limited understanding of directly stated information
- rarely connects relevant ideas and information to understand the meaning
- has difficulty integrating personal knowledge and experience to extend meaning

Complete

The student reads with reasonable accuracy and proficiency:

- demonstrates understanding of directly stated information
- usually connects relevant ideas and information to understand the meaning
- has moderate success in integrating personal knowledge and experience to extend meaning

Standard for Writing

Incomplete

The student has difficulty communicating in writing:

- demonstrates limited skill in developing, organizing and connecting main ideas with supporting information
- writes in a limited variety of forms and finds it difficult to complete such tasks as summary, opinion piece, news report and informational paragraph
- uses inappropriate or incorrect words and sentence structure
- makes errors in spelling, grammar and punctuation that impede meaning

Complete

The student communicates in writing:

- demonstrates an ability to develop, organize and connect main ideas with some supporting information
- writes in a variety of forms such as summary, opinion piece, news report, informational paragraph
- generally uses appropriate words and sentence structure
- generally does not make errors in spelling, grammar and punctuation that impede meaning

For further information on the standard-setting process, please see Appendix B.

Appendix C contains a report on a standard-validation process that took place in August 2000.

EQAO has also sponsored research to compare the Grade 10 test standard with the reading and writing standards used in the national School Achievement Indicators Program (SAIP) and two international assessments, Progress in International Reading Literacy Study (PIRLS) and Programme for International Student Assessment (PISA). The findings will be posted on the EQAO Web site in March 2001.

Applying the Pass/Fail Standard for Reading/Writing

This information will be posted on the EQAO Web site in March 2001.

12. Student Background Information

EQAO collected a variety of types of student background and test administration information that will be reported on the *Individual Student Report*. It includes the following:

- Level of study in Mathematics and English courses (academic, applied or locally developed)
- Accommodations or special provisions provided and the supporting details
- Exemption from the test in cases where students were not working towards a secondary school diploma
- Other reasons for non-participation in the test, e.g., student enrolled in ESL/ELD and not having sufficient language proficiency to participate, or student identified as exceptional, etc.
- Enrollment in an ESL or ELD program and the specific level of study
- Identification as an exceptional student by an Identification, Placement and Review Committee (IPRC) and the exceptionalities that apply. Please note that although data were gathered on the student's category of exceptionality, this information will not be reported on the *Individual Student Report*.

It will be used only in aggregated form for purposes of tracking and making recommendations related to the accommodation of exceptional students.

- The existence of an Individual Education Plan (IEP) for the student

For further information, please refer to Appendices A to F and to the EQAO Web site (www.eqao.com).

Appendices

Appendix A
An Overview of the Foundations for the Ontario Secondary School Literacy Test

Appendix B
Grade 10 Test of Reading and Writing — A Report on the Process by Robert J. Wilson

Appendix C
Validation of Standards, August 2000

Appendix D
Procedures for the Implementation of Accommodations for Students with Special Needs

Appendix E
Special Provisions for Students Enrolled in ESL/ELD Programs

Appendix F
Recommended Resources for the Ontario Secondary School Literacy Test

Appendix A

An Overview of the Foundations for the Ontario Secondary School Literacy Test

Assumptions about Reading

The contemporary view of reading acknowledges that reading is essentially a meaning-making activity. The reader is not simply using skills to decode meaning but is also actively engaged in applying personal knowledge and experience to understand, interpret and extend meaning.

This assumption about reading is embedded in the new Ontario elementary and secondary curricula. For example, in the elementary language curriculum, the focus on the importance of integrating knowledge and experience into the reading act is emphasized: "Reading is a complex process that involves understanding the relationship between written language and speech and relating ideas and information encountered in reading to one's store of knowledge and experience." In addition, one of the overall expectations across all grades states that students will use their own knowledge and experience to interpret what they have read. In the Grade 9/10 English curriculum, the importance of using knowledge and experience effectively as a reading strategy is emphasized. Both the Grade 9 applied and academic curricula in English, Geography, Mathematics and Science identify the requirement to relate, connect or apply information and ideas in texts to other situations, knowledge and experience. Connecting personal experience to understanding of texts is one component identified in subjects across the curriculum. As well, both elementary and Grade 9 curricula address the importance of the reader responding both critically and creatively to what is being read.

The design of the reading component of the EQAO Grade 10 test is consistent with the design of other contemporary assessments of reading, such as the national School Achievement Indicators Project (SAIP), which operates under the auspices of the Council of Ministers of Education, Canada (CMEC), and the Programme for International Student Assessment (PISA), which operates under the auspices of the Organisation for Economic Co-operation and Development. All three assessments are also consistent with the most recent theory, research and conclusions about the cognitive processes of reading.

The Grade 10 Literacy Test utilizes questions on two types of reading skills in response to the reading selections: those that measure students' understanding of explicit and some implicit information through the application of discrete reading skills; and those that measure students' interpretation, evaluation and extension of meaning through the integration of reading skills. The latter requires the students to filter their comprehension of, and response to, the material through their own knowledge, experience and reflection. Including such constructed-response questions acknowledges the key role that the reader has in the meaning-making process when reading a text.

Assumptions about Writing

Both the elementary language and Grade 9/10 English curricula acknowledge the complexity of the act of writing. The elementary language curriculum states that "writing is a complex process that involves a range of skills and tasks," and the Grade 9/10 English curriculum states that "writing is a complex process that involves a range of complementary thinking and composing skills." Both curricula suggest that writing is complex because it involves the integration and balancing of the three essential skills of composition: developing and organizing a range of ideas; selecting elements of style appropriate to purpose and audience; and applying the conventions of language and of form. The Grade 9/10 English curriculum, for example, states that "as writers compose, they consider their audience, make decisions about form, style, and organization and apply their knowledge of language use." Whether it is the curriculum for Canadian and World Studies, Health and Physical Education, Mathematics, Science or Technological Studies, the expectation is that students will communicate information and ideas using appropriate language conventions, for different audiences and purposes, in a variety of forms.

While all curricula emphasize the mastery of particular conventions of correctness as a key goal, they also indicate that students should primarily be taught writing in the context of authentic writing activities. The elementary language curriculum states that "basic conventions should be taught as much as possible through use and application in purposeful contexts. . . [by having students engage in] meaningful, creative activities." Similarly, the Grade 9/10 English curriculum states that "conventions are best learned in the context of meaningful and creative writing activities that allow students to develop the ability to think and write clearly and effectively."

The design of the writing component of the Grade 10 Literacy Test is consistent with the expectations for writing found across the Ontario curricula. To provide a range of opportunities for students to demonstrate the identified skills, the test requires four writing tasks on varying topics, for a variety of purposes and audiences in a variety of forms. The writing tasks were selected as representative of writing requirements from across compulsory subject areas; the writing tasks also reflect the types of writing needed in daily life. They have been designed to be manageable within the context of a timed test. Students were provided with advice on how to approach and complete each of the tasks on the test. This advice was provided in the *Support Materials* sent to the schools and available on the EQAO Web site. There is some expectation that students will develop, draft and revise their work to the degree possible within the time limits of the test. Assessment of writing through examination of completed on-demand student pieces is consistent with the design of other contemporary assessments of writing.

Essential Qualities in Reading and Writing

The design and development of the reading and writing components of the test began with the study of a number of sources, including theory and research, current Ontario curriculum policy documents for Grades 1 to 9 and the personal professional knowledge brought to the task by a development team whose collective and accumulative professional experience in language learning and teaching averaged 25 years per member.

Researchers and literacy specialists recommended establishing essential reading and writing skills to match the target population and the test mandate. As early as November 1998, it was determined by the Assessment Coordinating Committee that a blueprint should be based on "essential competencies" as required by the expectations of the curriculum.

The development team charted the types of reading and writing knowledge and applications required by the curriculum documents for language and the other compulsory subject areas for the skills and the forms or genres required by the expectations as well as the specific terminology used to describe the expected skills. The dominant elements were then identified. A tentative list of "essential qualities" was drawn up by the team and reviewed by members of the Technical Advisory Committee.

The Essential Qualities Identified for Reading (Precursors of the Reading Skills)

- Uses prior knowledge and/or personal experience to interpret texts
- Understands both explicit and inferred meanings
- Uses structural features of texts to clarify meaning
- Uses stylistic features of texts to clarify meaning
- Demonstrates understanding of a wide variety of literary and informational texts

The Essential Qualities Identified for Writing (Precursors of the Writing Skills)

- Develops and supports information and ideas in written work around a controlling idea
- Organizes and links information and ideas clearly and effectively
- Adapts written work to suit intended audience and purpose
- Demonstrates appropriate characteristics of a variety of written forms
- Applies conventions of standard Canadian English in written work

From the essential qualities, the specific reading and writing skills to be measured were identified and defined.

The Reading Skills

Key Skills

- Understands explicit meanings in a range of texts; for example, locating explicit information, making links between explicit statements in the text
- Understands inferred meanings in a variety of texts; for example, supporting inferences with reference to text, generalizing or concluding from explicit and implicit information in the text, making connections between implicit ideas and information in the text
- Interprets and understands meaning in texts by integrating textual information with reader's own knowledge and experience; in other words, the interpretation or extension combines information and ideas in the text with information and ideas that the reader brings to the text; for example, responding to both affective and/or cognitive domains; making judgements about tone, style and purpose; making analytic extensions

Supporting Skills

- Uses knowledge or context to understand familiar and unfamiliar vocabulary and technical or imaginative uses of language required to understand meanings in texts; for example, word choice, connotation, idiomatic expression, colloquial language, slang, dialect, jargon, imagery, metaphoric expression, similes, personification, puns

- Uses grammatical and syntactical structures as well as punctuation to understand meanings in texts; for example, subject, predicate, subordination, sentence type, tense, agreement, antecedents; parallel structure, repetition, juxtaposition, rhetorical questions, sentence variety, rhythm, alliteration; comma, period, exclamation/quotation marks, colon, dash, italics
- Uses knowledge of overall structure and organizational elements and patterns to understand meanings and purposes of texts; for example, paragraph structure, topic sentences, introduction/body/closing, transition or connecting words and phrases, paragraphing, parallel phrasing; plot, character, setting, conflict, point of view, foreshadowing, mood/atmosphere; examples, compare/contrast, cause/effect, sequence
- Uses knowledge of visual or extra-textual features to understand meanings and purposes in text; for example, titles, illustrations, diagrams, charts, icons, captions, margin notes, columns, shading, fonts, font size, bolding, underlining and layout

The Writing Skills

- Demonstrates the knowledge and skill needed to develop a controlling idea with supporting details
- Demonstrates the knowledge and skill needed to organize and link information and ideas
- Demonstrates language awareness through word choice and sentence structure
- Demonstrates knowledge and skills needed to produce written work in a variety of forms to suit intended audience and purpose
- Demonstrates knowledge of written conventions through correct use of grammar, usage, punctuation and spelling

The Design of the Test Instrument

The five-hour test is divided between reading and writing tasks and interspersed over two days. Each student is required to read 12 reading selections, answer 100 questions on them, and complete four pieces of writing.

The design of the test is determined by a table of specifications and a blueprint. The table of specifications determines the relative number of reading questions by type, skill and level of difficulty as well as the four types of writing tasks. The blueprint describes the number and sequence of reading selections, the number and types of questions attached to each reading selection and the placement of the writing tasks.

The Reading Component

The test uses three types of reading selections, including the following:

- Informational (50%), such as explanation, instructions
- Graphic (25%), such as schedules, graphs, tables
- Literary (25%), such as story, description, dialogue

Informational texts are written to communicate information, including explanations, opinions, presentations of technical or research data. Graphic texts are informational texts in which charts, illustrations and other forms of visual layout are used to communicate the information. Literary texts are those types of texts that represent imaginative expression in the writing arts. These texts tend to be read for leisure or pleasure rather than strictly to obtain information.

In general, to ensure a wide variety of texts reflective of the curriculum expectations and real-life reading, the test contains 12 texts of varying lengths. Shorter texts tend to be more accessible to weaker readers. In order to be authentic to reading practice and to challenge all levels of reading ability, the test has texts ranging from 150 words to 1200.

The selection of texts for the test has been based on specific criteria beyond the genre and qualities required by the table of specifications. These included a search for informational and literary selections that

- appeal to the interests of the target audience;
- respect the needs of Ontario's diverse student population in terms of geography, rural and urban environments, heritage and values;
- vary in length but do not exceed 1200 words;
- deal with information, ideas, themes and/or issues from a global or Canadian perspective;
- require familiarity with forms, organizational patterns and language use corresponding to the expectations in *The Ontario Curriculum* for the target audience;
- do not depend on detailed specialized subject knowledge whether or not it is required by *The Ontario Curriculum*;
- have a level of language, abstract thought and complexity of structure appropriate for the target audience; and
- attempt to exclude material from textbooks or anthologies currently studied in Ontario schools.

The texts were also selected for three levels of reading difficulty with the largest proportion of texts having a medium level of difficulty. The assessment of difficulty is based on the following criteria: familiarity of subject matter, level of language, level of abstraction, complexity of structure, use of stylistic devices and length.

Educators and specialists in language and literature were asked to select passages for validation with attention to age/grade appropriateness, equity or bias, links to the curriculum, literary quality, accessibility and relevance over time.

Reading selections represent a point of view influenced by personal and social identities, values and experiences. It is essential that the reading selections on the test represent balanced and inclusive perspectives. Validation Committee members, with expertise on social and ethnocultural issues, special education, equity and language, were asked to examine the validity of the selections in terms of age/grade appropriate vocabulary, age/grade appropriate content, gender bias, socio-economic bias, ethnocultural and racial bias, religious and values bias, disabilities bias, and relevance over time. While hundreds of texts were selected for consideration, the validation process reduced the total for use to about 10% of those considered.

The test uses three types of items (questions):

- Multiple Choice (40%)
- Constructed Response (35%), requiring short answers
- Constructed Response with Explanation (25%), requiring short answers in which students justify or explain their thinking

Each test item is designed with one of three levels of difficulty. The level of difficulty is first anticipated, then piloted and field tested for confirmation. The factors determining the level of difficulty include the following:

- type, quality and quantity of information required
- scope of the thinking required (number and intricacy of steps, abstraction level, familiarity)
- complexity of information to be integrated or number of relationships to be made
- the extent to which the focus is on what is said or on how it is said
- complexity of the item wording
- the plausibility of the options (in the case of multiple-choice items)

The Writing Component

To provide a range of opportunities to demonstrate the writing skills, the test requires four writing tasks. The tasks are representative of writing expectations across the compulsory subject areas and in daily life. The writing tasks are divided equally across the two test booklets to be completed along with the reading tasks over two half days. Students are asked to write the following pieces:

- a summary of an opinion text in a single paragraph
- a series of paragraphs expressing an opinion
- a newspaper report based upon a topic/photograph provided
- a paragraph based upon information provided

Each task requires participants to demonstrate all six writing skills. Three of the tasks are considered to be short in length, a page or less. The series of paragraphs expressing an opinion constitutes the longer writing task. There is a relationship between the reading component and the writing component of the test. The series of paragraphs expressing an opinion, the news story and the information paragraph are preceded by models of these genres in the reading component. The summary requires students to summarize one of the reading selections.

The *Support Materials* sent to schools and available on the EQAO Web site illustrate the expectations and requirements of each writing task. The *Support Materials* also describe for students the exact instructions, language, format and space for writing on the test.

Students are required to comply with the specific instructions for each task in the blueprint. The intent of the writing component is not to measure the general writing ability of the student but to measure the student's skills in writing for a particular purpose and audience in a particular form, as required by the curriculum expectations in Grades 7 to 9. The audience in each case is an adult seeking information or opinion. On-demand writing allows only a limited time for a writing process. Dictionaries, thesauri, peer or teacher assistance are not available to the students.

The Specific Characteristics of the Four Writing Tasks

Summary (approximately 20 minutes)

The students are given a short text of approximately 200 words. They are asked to write a summary of about 50 to 100 words in which they express a main idea and at least two supporting details (ideas) from the text. The summary is a common requirement in school subjects such as geography, science and technology. Students will have encountered the text as part of the reading component where the questions provide a preliminary examination of the content of the text.

Series of Paragraphs Expressing an Opinion (approximately 40 minutes)

The students are given an issue familiar to their age group and asked to develop a series of paragraphs in which they express an opinion on the issue. The students are required to write a minimum of three paragraphs with an introduction, development and a conclusion. The series of paragraphs is integral to most secondary school courses as well as in the course of daily life when people articulate concerns to businesses, newspapers, public institutions and community groups.

News Report (approximately 20 minutes)

Students are given a photograph and a headline. They are asked to develop a newspaper-style report. This task reflects the significance of media in contemporary daily life and the fact that it is one of the four strands in the Grade 9 English curriculum.

Information Paragraph (approximately 20 minutes)

Students are provided with a list of information relating to a topic. They are asked to develop a paragraph, with a single topic sentence and supporting details drawn from the list provided. This task reflects research and report writing common to most subject curricula.

Appendix B

The Grade 10 Test of Reading and Writing — A Report on the Process by Robert J. Wilson

Introduction

Standard-setting is not an activity most of us engage in on a regular basis. In most real-life situations, standard-setting has already occurred and we are faced only with the question of whether or not the standard has been met. Driver's tests are a good example. The examiner has a checklist, points have been assigned to various aspects of the test (a low cost for failing to adjust the mirrors; a higher one for driving through a stop sign), and we are delighted if we meet the standard and chagrined if we do not. The standard itself is not at issue. Presumably, there has been some valid relationship found between high scores on the test and an ability to drive safely.

In achievement testing in education, however, standards are intimately connected with the particular test that is used. The measures found in the driver's test, for instance, are fairly objective, but constructors of learning assessments intentionally or unintentionally shape the scores by selecting which items appear on the test: easy items make the scores higher; more difficult ones make the scores lower. Standards in education, then, cannot be validly arrived at without some understanding of what the test itself required students to do.

This shifting of difficulty across achievement tests also means that the score by itself is not a good indicator of the competence examinees have shown. Perhaps one student has answered all the easy questions correctly — is able to recall discrete facts she or he has memorized — but none of the more sophisticated ones. Another student misses many of the easy questions but shows much greater ability at putting together complex information. In the driver's test example, this issue is handled by making some items more critical to pass than others.

Sliding over these truisms, we sometimes conclude, for example, that "50% is a pass" but clearly that judgement leaves unanswered two very important questions:

- What does 50% mean on this particular test?
- What does a pass tell about what the student can do?

Questions such as these give rise to the need for standard-setting.

Definition of Standard-Setting

Standard-setting procedures describe in words what various levels of performance mean, and then link those levels with actual displays of that performance. Descriptions of standards arise from what both standard-setters and test constructors value. The test constructors build their values into the choice of questions they set for the examinees. The standard-setters characterize those values and relate them to actual achievement.

It is obvious, then, that people's beliefs form a large part of any standard-setting practice. At root, it is a subjective process, like that used by juries in arriving at a judgement of guilt or innocence. What is vital in all these kinds of judgements are defensible procedures and an assessment in which those procedures were actually followed.

Approaches to Standard-Setting[1]

There have been three decades of research and development into standard-setting for achievement tests, much of this work having been conducted in the United States. The early work developed when the multiple-choice item was used almost exclusively in high-stakes tests and an atomistic view of learning was paramount.

One popular method, usually called the Angoff after the person who first suggested it, exhibits both these traits. Standard-setters are presented with multiple-choice items and asked to consider how "minimally competent" individuals would do on each question. The ratings standard-setters give are frequently subjected to comparisons with those of other standard-setters, and an iterative process follows until some agreement among raters can be reached. The ratings are then summed, and the passing score for a test decided.

Many variations of this approach have been adopted. In Canada, one of these is currently used with the School Achievement Indicators Project, conducted by the Council of Ministers of Education.

Along with other "test-centred" models, the Angoff approach makes certain assumptions. Chief among them is that items are discrete, and can be assessed independently of one another. Second, the model assumes that a simple sum of the item probabilities constitutes a score that is sufficient to meet the standard. More modern assessments, like the Grade 10 Assessment of Reading and Writing, do not rely exclusively on discrete items, especially where it becomes obvious that some items are more important than others in the overall assessment.

Another type of standard-setting depends on judging the examinees rather than the items. Judging item difficulties, it is maintained by proponents of this other approach, is too difficult, especially when judgements are made by those unfamiliar with schooling and students on a day-to-day basis. It is more reasonable to consider a person competent or incompetent in some area rather than performing the more abstract task of estimating probabilities of items.

One of the more common versions of the "examinee-centred" models is called the contrasting groups method. In this approach two groups of examinees are created by definition: one as competent and the other, not. When their performances are graphed, a point or score is located that maximizes the probability of making a correct judgement about those who have performed well compared to those who have not. One of the issues with this method is the degree to which the original classification is valid. A variation on this approach is to use a borderline group, rather than two separate groups, which avoids the problem of definition but creates a higher need for precision in judgement.

While there are competing models and competing conceptions of what makes valid standard-setting for performance assessments, there is some general consensus about what is required of any approach to this task. These are summarized below:

- Standard-setting approaches need to be conducted with a view to providing evidence for validity issues. This evidence will not only be empirical but also judgemental. The procedures must be defendable as reasonable and judicious.

1 The general approaches to standard-setting are explored further in Cizek, G.J. (1996). Setting passing scores. *Educational Measurement: Issues and Practices, 15* (2), 20–31. A good review of standard-setting in performance assessments is available in Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research, 64,* 3, 425–461. Each of these articles also has an extensive bibliography of other relevant studies.

- The test itself should provide evidence of its own validity, especially in terms of content and construct adequacy.
- Descriptions of the performance standards should relate both to the assessment and to the uses to which the results will be put. In essence, this means that the test should measure something worthwhile and the standards should express those values in their descriptions.
- Documentation should be available that describes the standard-setting, specifically, how they were selected and trained. There should also be careful documentation about the materials that were developed and used in the process.
- Reports on the implementation of the process should be prepared and made available to interested parties.
- Records should be kept on the participants' perceptions to determine if their understanding of the process matched those of the designers.
- Subsequent evidence should be collected that supports the reasonableness and appropriateness of the standards and their applications.

The next section documents how the Grade 10 standard-setting exercise attended to these criteria.

The Nature of the EQAO Process

Expert Review

The overall structure of the plan for the Grade 10 Test of Reading and Writing standard-setting was sent to three experts in the field for analysis. The reviews resulted in much discussion of key issues that needed to be addressed. Many of the comments were incorporated into the plan; some were not. The reviewers did not see the final product that ensued, so their comments were of necessity limited to the broader design.

Participants

The first criterion applied to the selection of the standard-setting participants was that there should be equal representation of members from the education profession and members from the public. A second criterion was that participants were representing themselves and not any groups to which they might belong. A third criterion held that the diversity of the province be represented insofar as that was possible with 18 people.

Directors of Education nominated educators they believed represented excellence in education. The public members were nominated by both educational and community representatives on the basis of experience with students or graduates of the public systems, especially in those areas where reading and writing skills were consistently applied.

The non-educators on the panel came from western as well as eastern Ontario, the mid-north (Sudbury and North Bay) as well as farther north (Thunder Bay), and from Toronto and surroundings. Three were business people, two worked with social agencies. There was also a homemaker, an arts administrator, a broadcast journalist and a community college department head. The public education group was made up of five teachers, a consultant and three administrators with approximately equal representation from Catholic and public schools. They, too, represented the geographical diversity of the province.

These nine educators and nine non-educators came together on the morning of May 30, 2000, and worked for two and a half days, including evening sessions. Other individuals participating in the process were there as facilitators. Three educators with experience in working in teams acted as

small-group leaders. Staff and consultants from EQAO implemented the arrangements, monitored the process, provided expertise in the content being assessed and generally oversaw the work. Another consultant with experience in standard-setting in British Columbia and on international assessments designed and implemented the specific tasks that constituted the agreed-upon process. Others acted as formative assessors and aides to the procedures as these were applied.

Materials

Background on the Grade 10 Test of Reading and Writing Skills

An important contextual background to the testing program was provided in both written and oral form. Participants were informed that the test would be given to Grade 10 students who are two and a half years away from high school graduation. It is not, therefore, a test of readiness for college, university or the workplace. The average age of the students taking the test will be 15.

The testing program itself fits within a wider course of educational reform in Ontario, where length of time to graduate has been changed from five years to four, where a new curriculum has just been built and where new emphases have been placed on career education and teachers as advisors.

The test is put together by staff from the Educational Quality and Accountability Office, the arm's-length agency of the government. This office is also responsible for the setting of standards. As devised, students take the Grade 10 test over five hours, spread over two consecutive mornings. It is based on generic skills of reading and writing found in several subject locations in the provincial curriculum.

Various policies concerning deferrals, accommodations and administration were also provided to the panelists. Exemptions, for example, will be narrowly defined but provisions will be developed for students with special needs. Some students could have the test deferred under certain conditions. Re-takes will be allowed, and schools will be required by the Ministry of Education to provide remedial help for those who fail.

Test Content

The test contains two sub-tests: one in reading and one in writing. Scores for each sub-test will be reported separately. In reading, students are asked to read materials that are informational (e.g., an article from a newspaper), graphical (e.g., a train schedule) and literary (e.g., editorial from a local newspaper). Questions on those materials could take one of three forms: multiple-choice, short answer (called constructed response) and more extended responses (called constructed response with explanation). The skills that were most emphasized in the reading portion of the test were those involved in understanding explicitly stated ideas, making inferences on the materials read and interpreting what was read in personal terms.

The writing portion of the test requires students to produce a summary of a passage they have read, a news report, an informational paragraph and a multiple-paragraph essay expressing and supporting an opinion. The goals of the test are to determine how well students are able to develop a main idea in their writing, provide supporting details and organize and link information. All of these are to be accomplished using appropriate words, sentence structure, grammar, spelling and punctuation, and should illustrate an awareness of audience, form and purpose. These aims were summarized into an assessment of the written production according to its content, form and organization. Tone was also assessed and the conventions of writing (spelling, punctuation and syntax) were examined throughout.

Test Materials

The materials most essential to the process of standard-setting were the tests themselves, along with examples of student work. While there are several forms of the tests, one complete form was used throughout the exercise as representative of the others. (When the tests are actually used, the scores from any form will be made comparable to those that would have been obtained from any other form.)

Twenty papers were selected for examination by the panel. Sixteen of these papers were chosen to represent achievement that might span a range separating complete from incomplete performance. Four others represented performance that reflected excellent responses to the task. The main objective in selecting the test materials was to provide the committee with examples of student work that covered the range of performance obtained during the field trials of the forms. Panelists were informed of the range of responses but were given no expectations at all as to how many students might be expected to, or actually did, achieve on the forms used in the standard-setting.

Guidelines for Writing Standards

Another important resource for the panel was the guidance given on the actual writing of standards. They were told that standards, to be most helpful, should be descriptive of what students at a particular level actually do (as opposed to what they cannot do, or do poorly). To be valid, standards needed to be referenced to the actual skill areas in reading and writing that the tests measure. Important parts of any standard are the levels — the descriptions of the various degrees of the standard. Thus, standards need to be clearly stated, sequential in nature and moving from descriptions of incomplete, through to complete, and finally, to exceptional work along a clearly understood continuum.

Stages in the Process

Reading

When the 18 panelists met for the first time, they were provided with much of the background material described above. After introductions, they were encouraged to ask questions and inform themselves as much as possible about their roles and goals. The leader of the standard-setting process then provided an overview of the entire process and objectives for the meeting. A discussion of the skill areas in reading and writing then ensued, followed by a review of answers and scoring used on the tasks themselves.

Panelists had previously been randomly assigned to one of three groups according to background (educator or public). The groups, functioning for the first time with their group leaders, were given two tasks. The first involved separating student responses from 20 papers relating to the skill area of gaining explicit knowledge from reading into three categories: incomplete, complete and advanced. A brief specification of what these terms might mean was provided. Incomplete, for example, was described as a performance that did not successfully meet expectations. Other than these largely tautological statements, panelists were urged to apply their own judgements to the papers. Even at this early stage, some panelists began collecting words they believed characterized a standard they would apply to each category.

Each individual did his or her sorting first. Then, within the group, a discussion took place to determine where the group would place each paper. Groups attempted to come to a consensus about the placement of the papers if they were able to; if not, they were to record where individual

decisions placed the papers. Those papers for which there was no consensus constituted a fourth category of papers: "not agreed."

Groups reported that, during this first task, some effort had to be made to deal with issues individuals presented. For example, criticism of individual test items had to be resolved if panelists did not believe a question was valid. Groups were encouraged to concentrate on the student responses, rather than the items, in their considerations. One group leader timed the discussions in her group and found that more than half the time was spent on the objectives of the group work, a third on reflection on the task, and one sixth each on interpreting the group's efforts and phrasing a response.

In reading, the process was repeated two more times: once for implicit meanings gained from the reading, and once for interpretation of texts. As the process was repeated, panelists reported more familiarity with the process and were glad to have started with the explicit meaning as it was generally seen as simpler than the other two.

Once all three skills sets were completed, the small groups gathered in a plenary session. The results for all three groups had been collated and were then presented on overheads so that all members could see what the other groups had done. Papers on which there was total agreement were set aside. Those where the ratings consisted of two groups' agreeing along with a majority of panelists were then classified with the majority. For papers where agreement was not reached, the groups reconvened to determine whether any changes needed to be made to their judgements. Decisions were then communicated to the larger group again, and a final decision was made following the rules described above.

The product of this stage of the exercise was a large group judgement about which papers constituted incomplete, complete and advanced work for the three major skill areas assessed by the reading test.

Standard Development

At this point, the small groups reconvened and were provided with a different task. Beginning with a description of what a standard is, individuals were asked to provide descriptors for the three reading skills that characterized the three groups of papers. For this stage of the process, the group leaders were absent. The detailed work of deciding on a standard description was accomplished by the panel members alone.

One of the difficulties with this part of the process for some panelists was a perceived lack of knowledge of what the expectations were for students in the area of reading, combined with a lack of experience with writing standards. Nevertheless, each group came up with descriptions as requested. A small group of facilitators, including a professional writer and a subject-matter specialist, then attempted to summarize the three groups' descriptions into a common statement. This common statement was then presented to the large group. This group made edits to the standard and then voted on the wording of the final statement.

Writing

Much the same process was followed for the writing portion of the test. There were some differences, however, in that writing was not divided into sub-skill sections. The examinees had produced writing samples, and these writing samples were scored by characteristics. Most characteristics were evident in all samples. Because of the nature of this assessment, there was no need to combine the parts of the test into one whole test as was done for reading. Each panelist

read each student's entire product. These differences, as well as information about the tasks and the scoring, were presented to the panelists in a large-group setting by a subject-matter expert from EQAO.

Following this orientation, panelists repeated the process of using individual judgements to arrive at incomplete, complete and advanced judgements of the writing papers. Then, working within their small groups, they arrived at a group judgement about the classification of these papers.

With the application of the large-group rules, all the writing papers were subsequently categorized into the three categories, and the description of standards ensued. These statements were also collated and summarized by a team from EQAO, and the summary evaluated by the panelists in a large-group setting. This statement was edited, agreed to and became the standard for writing.

Sorting into Complete and Incomplete Categories

The final work of the panelists involved taking the 20 papers in reading and writing and categorizing them as complete (pass) or incomplete (fail). As this is the level of the decision that will ultimately be communicated to students, no attempt was made to make any finer distinctions.

Following the pattern already well established, each panelist made his or her judgement first about whether a paper met the standard or did not. The small groups then attempted to come to a consensus on whether or not they agreed that the paper was complete or not. Finally, the large group resolved any differences that remained using the criteria established before.

How that process worked is what constitutes the next section of this report.

Results

Analysis of the Process and Products

To review, the process of standard-setting involved four steps. First, each individual panel member read twenty papers in reading (divided into three separate sub-tasks and then summarized as a whole task) and repeated the process for writing (without the sub-task element). Second, each group of six panelists discussed each of the 20 papers at each of these steps and attempted to arrive at a group consensus as to whether the papers were incomplete, complete or advanced. In a small minority of papers, consensus was not possible.

Third, each member reviewed the results of all the groups in a large-group session and then returned to the smaller groups to see if any adjustments needed to be made. The final decision about each paper was reached when at least two of the three groups agreed on the designation and a majority of individuals agreed. Papers falling within each of these designations were then analyzed within the small groups and then again in the large group to come up with a description of complete, incomplete or advanced, first in reading and then in writing.

Finally, when those descriptions were agreed upon, each individual then decided whether the 20 complete papers in reading and then again in writing warranted a complete (pass) or an incomplete (fail) designation. Small-group and large-group consensus was then sought on these categorizations as well.

Individual Ratings of Reading and Writing

When panelists were asked to judge the 20 papers in reading, one panelist failed only one paper while another failed nine. In writing, two panelists found only one paper lacking while another

judge tagged fourteen papers that way. This large range shows that individuals did bring their own criteria to bear on the process initially and did their first sorting of papers with those expectations in mind.

Background of Panelists

Because half the individuals involved were chosen from education backgrounds, while the other half represented the broader community, the question arises whether judgements tended to reflect the judges' background. Were the public members, for example, more severe than their colleagues from education? With one exception, they were not.

Small-Group Ratings

What happened to individual ratings when these were discussed within the small-group setting? In all three groups, the variability between individuals became much less as each of the three small groups attempted to arrive at a consensus to present to the larger group. The discussion in the small groups tended to produce a judgement about student performance that benefited from the disparate opinions of the individuals. It is also evident that for that consensus to occur, every voice was heard and respected.

Large-Group Ratings

This same pattern of individual ratings being significantly related to the group recurs as the analysis moves to a consideration of the large group. In this analysis, there is only one instance of an individual's rating failing to correlate with the final large-group rating in reading and in writing. Further investigation reveals that this individual's ratings in writing were unrelated to those of most other members. It follows from this that each group's ratings were also significantly correlated with the final rating given to reading and writing, except in the one group to which the outlying member belonged.

In the end, 40 judgements had to be made about the passing or failing of twenty papers in reading and in writing. Thirteen of the judgements in reading and eight of the judgements in writing were unanimous across the groups. After group reflection, an additional four judgements in reading and six judgements in writing were agreed upon by all. Finally, three judgements in reading and six judgements in writing had to be made with the majority rule. Thus, 31 (78%) of the decisions were made unanimously across the groups and nine (22%) were subject to majority vote.

Summary

What these data show is that individuals began the process with divergent ideas about what would constitute acceptable performance on these tests. When discussion around their respective points of view occurred in the small-group setting, many of the differences became subsumed under a group consensus. When this small-group consensus was taken to the larger-group setting, each group's views contributed to a final judgement in which virtually all individuals as well as groups correlated to a significant degree.

Reactions to the Process

A 16-item, Likert-type scale was administered to the panelists at the end of their work. Space was also provided for commentary. Frequencies, correlations and factor analyses were applied to the ratings, and content analysis was performed on the comments.

Several of the items asked respondents for an opinion about the process itself:

- Were the instructions clear?
- How did their group function?
- Was enough time allowed?
- How did the mix of the public with educators work?

On all of these questions, the panelists responded very positively. Their comments reinforced the conclusions from the quantitative data:

- "Instructions were made clearer as process developed. First morning was most difficult."
- "Straightforward — to the point"
- "Well Done! Both in small and large groups"

Some individuals felt a pressure of time on all decisions. A few individuals experienced difficulty with the communication before the process. In the end, the vast majority of the panelists were confident of the standards they had reached in both reading and writing.

A second grouping of items centred on the value of the process. Panelists found it rigorous but effective and would support its use again in similar circumstances. One comment that summarized this general feeling was as follows:

- "This was a very productive and rewarding experience. Thank you for the opportunity to make a contribution."

The tasks however were not seen as trivial. Panelists took their roles seriously:

- "I found the application more difficult than I had expected. It felt so final."

Another reaction of the group concerned the tasks they were assigned. There was massive support for the clarity of the directions they were given in the large-group setting, the setting where the most important and final decisions were made. Some comments speak to the nervousness some felt towards what they were being asked to do:

- "Although at times we were learning at the same time, this wasn't necessarily a bad thing."
- "We all had different views but only one could be accepted."

When asked directly about the process, all but one panelist either agreed or strongly agreed that this approach to standard-setting was useful and effective in meeting the goals.

Summary and Next Steps

Standard-setting of some kind is an inevitable consequence of implementing high-stakes testing. The model that was used for the Grade 10 Test of Reading and Writing was deliberately different from that used elsewhere. Here, an equal number of members of the public and educators were engaged in all facets of the exercise. They were involved first and foremost in actually characterizing the standards that would be used. Their description of a standard for reading performance and another for writing performance included the distinctions among the levels that could be utilized in characterizing students' performances.

The panel arrived at these descriptions through an examination of the test booklets and student work on the tests. The panel had to keep in mind the characteristics of the Grade 10 student population in Ontario, where students were in their schooling history and where the test fit within the context of school reform. Panelists brought a unique perspective to the task as evidenced by their diverse opening opinions. Working through the standard-setting tasks, first with other members in small groups and then together in a large group, these 18 people arrived at a formulation for acceptable work on this test for both reading and writing.

After creating the categories for both subjects, they were then able to categorize 20 booklets of student work into those that met the standards and those that did not.

At no time did the panelists consult a distribution of scores of students on these tests to temper their judgements. They observed samples of student work that spanned the achievement range on the field trial instruments but with no understanding of how many students might be expected to fit any one of their criteria. In this important respect, this standard-setting exercise followed that which is used in other significant areas of our life where standards refer to goals that the community expects should be both achieved and achievable.

Next Steps

While the most significant work is completed, there remains the establishment of a cut score to complete the task. When future students take the test, they will be judged to have passed or failed on their overall scores on each of the reading and writing forms.

Steps will have been taken to ensure that the various forms of the test are equitable. For each test, a cut score will have to be established that separates acceptable from unacceptable performance. To make the derivation of that score as reliable as possible, it is recommended that many more than the 20 papers be used than these panelists examined as part of their work.

Another group needs to be established — call it a standards-application group — that would take a larger group of papers and perform the last part of the original panel's work; namely, use the standards to categorize papers into complete and incomplete categories. The papers used by the original standard-setting group could be made part of the larger group of papers.

Several important things would be accomplished by this exercise. First, a reliability check on the first panel's work would be performed. The question to be answered is, "To what extent would another group using the first group's standards arrive at the same judgement as to whether the examined papers were acceptable or not?" This check is important if we are to maintain that the original decisions were not unique to this particular group of raters but represent more generalizable conclusions. Second, the establishment of a cut score needs to be as precise as possible. Twenty papers are too few to have this important function performed accurately. If one or two of the papers are aberrant for whatever reason, the whole sample and the conclusions from it could be distorted. With more papers involved, the impact of any odd performances by a minority of students would be minimized.

Finally, when the standards are actually applied to real student work it is recommended that student papers whose scores are lower than the cut score but within a margin of error be marked a second time by a trained marker. If the second marker agrees with the judgement of the first one, the student receives an "incomplete" rating. If there is a disagreement, the paper will be marked again, and the majority conclusion communicated to the student.

This paper was written just after the standard-setting activity in May 2000, therefore, prior to the first administration of the Grade 10 test. The name of the test has been changed from the Grade 10 Test of Reading and Writing Skills *to the* Ontario Secondary School Literacy Test.

Appendix C

Validation of Standards, August 2000

The Purpose

The purpose of the second session of the standard-setting process was to validate the standards established by an 18-member committee that met in May 2000.

The validation committee of 20 educators and nine members of the public was divided into teams of five, with a balanced distribution of gender and professional roles. The participants were to classify each student booklet in reading and each set of four tasks in writing as "unacceptable," "acceptable" or "advanced" based on their examination of the whole performance and their understanding of the standards written in the first session. Each of the 100 reading and 100 writing booklets were to be classified a minimum of five times. Included in these booklets were the 40 booklets used in the May standard-setting process. This would ensure that validity and comparability could be assured.

The Process for Reading

Participants were provided with background on the details of analytic reading and trained through participation in a mini-version of the original standard-setting process. Participants were also presented with information about the test design, including details on the three key skills, the types of questions, the levels of difficulty of the questions and background on compliance issues.

Participants were asked to use the "Scoring Guide for Test Questions, Skill 1, Explicit" to discuss sample papers as unacceptable, acceptable and advanced papers.

They were then asked to practise applying their knowledge of the analytic scoring guide and their understanding of the sample papers to five booklets containing only the Skill 1, Explicit responses. They discussed their classifications with group members and attempted to reach some consensus. They were told that consensus was not essential since this was only analytic training to familiarize them with the detailed knowledge of the test and the acceptable responses.

The results of the practise session, including both the number of individual choices in each classification and the consensus reached, were then presented to the committee as a whole. The same procedure was then used to examine Skill 2, Implicit, and Skill 3, Interpretation. Participants were then presented with the statements of standards written during the May standard-setting process. They examined how these statements embraced the three key skills and the careful distinctions in descriptors between unacceptable, acceptable and advanced.

They were asked to review the scoring guide as a whole and to begin studying the sample papers provided for each classification (unacceptable, acceptable and advanced). They discussed as a team why each sample represented the category assigned. They were then asked to classify the five papers provided. Reliability papers were used to determine readiness to apply the standards on 100 papers and to ensure reliability throughout the process. As soon as a booklet was classified five or six times it was withdrawn.

The Process for Writing

Participants were asked to look at each piece of writing holistically. They were asked to talk about writing in terms of the skills, study how the skills were embraced in the standards and then apply the standards to sets of student writing.

Participants were provided with background in the following areas: writing skills, contextual demands for writing, the four specific writing tasks (summary, opinion, report and information paragraph) and the key assumptions about quality and compliance applied by scorers.

Training papers and reliability papers were used to establish readiness to apply standards to the 100 writing papers. The statements of standards were then presented and examined for the presence of the skills, the careful wording of which distinguished acceptable, unacceptable and advanced. Participants were asked to examine each of five anchor papers, two for unacceptable, two for acceptable and one for advanced. Each anchor paper was discussed in light of the statements of the standards. Participants were told that the anchors were highly significant in writing because there was no scoring guide. There were only the standards and the samples to ground their judgement.

Participants then practised on four sample sets of writing followed by a discussion by the committee as a whole. After this general discussion, participants began to classify the sets of writing. Two reliability papers were used to ensure the consistency of the group in their work. All papers were classified at least five times.

Appendix D

Procedures for the Implementation of Accommodations for Students with Special Needs

I Introduction

This document outlines the procedures for implementing the "Accommodations" section of the Ministry of Education Policy/Program Memorandum on Accommodations, Deferrals and Exemptions.

II Definition of Terms

STUDENTS WITH SPECIAL NEEDS

"Students with special needs" are defined as students with an Individual Education Plan (IEP) who **may or may not** have been identified as exceptional through the Identification, Placement and Review Committee (IPRC) process.

ACCOMMODATIONS

Accommodations are supports and services that enable students with special needs to demonstrate their competencies in the skills being measured by the test.

Accommodations change the way in which the test is administered or responded to by the person being tested, without altering what the test is designed to measure and the meaning and validity of the test results.

MODIFICATIONS

The term "modifications" refers to those supports and services that are different from accommodations because they alter the meaning and validity of the test results. These supports and services include adaptations/alterations of the test content and teacher assistance in the form of clarification or explanation of parts of the test.

Notes on the distinction between "accommodations" and "modifications" and issues arising from that distinction:

A province-wide review of IEPs is currently underway. This review will ensure consistency in terminology across the province, including consistency in the definition of "accommodation."

Because local review processes may not be complete, consistency of IEPs cannot be ensured for the October 2000 administration of the test. What is included in IEPs in October 2000 may not always conform to the definitions in Ministry of Education Policy/Program Memorandum for "accommodations" and "modifications."

Schools must recognize that modifications, as defined in the Policy/Program Memorandum, are not allowed, even if they are called "accommodations" in IEPs in October 2000.

The nature of the test and the conditions imposed on its administration by the standards of practice for large-scale testing are such that certain strategies that are appropriate in a classroom

setting are in fact modifications as defined in the Policy/Program Memorandum and are not allowable for the purposes of this test.

Examples of strategies that are appropriate in a classroom instruction of assessment context but that are not allowable for the Grade 10 Test of Reading and Writing Skills are clarifying questions for students, encouraging students to question for clarification and additional information, and encouraging oral discussion prior to writing tasks.

In order to ensure that all students have opportunities to familiarize themselves with the test and receive clarification and explanation of the types of questions and writing prompts and of the language and layout of the test, EQAO has posted *Support Materials* on its Web site (www.eqao.com). Please see Part VII, "Familiarizing Students with Test Format and Instructions Prior to the Test" for further information on these materials and their use.

III Guiding Principles

All students working towards a diploma, including students with special needs, should have the opportunity to participate in the Grade 10 Test of Reading and Writing Skills in order to meet the requirements for an Ontario Secondary School Diploma.

Accommodations allowed by Policy/Program Memorandum and these "Procedures for the Implementation of Accommodations for Students with Special Needs" will not compromise the integrity or validity of the test.

Schools must not avoid accommodations because providing them is sometimes challenging. The *Ontario Human Rights Code*, 1981, and other legislation pertinent to educational settings set out "duty to accommodate" provisions for all persons with disabilities.

Adherence to the policy and procedures for accommodations, and careful documentation of practices will ensure consistency and fairness across the province and from one year to the next.

IV Procedures for Accommodations

Accommodations must be made for students with special needs. (See definitions in Part II.) Students qualifying for accommodations **may or may not have** been formally identified as exceptional by an Identification, Placement and Review Committee (IPRC).

Decisions about accommodations

Decisions about accommodations must

- always be made on an individual student basis;
- be made by the principal in consultation with the student, parent, appropriate staff (including Special Education resource services staff, where appropriate);
- be made prior to the assessment;
- take into consideration all accommodations included in the student's IEP, which are also listed for use on the test;
- comply with the Grade 10 Test accommodations policy and procedures as outlined in Policy/Program Memorandum and these "Procedures for the Implementation of Accommodations for Students with Special Needs";
- be clearly communicated to the student and parent/guardian in advance of the writing of the test; and
- be documented in full on the *Student Information Form.*

Procedures for resolution of disagreements

There are two categories of requests for consideration of accommodations over which there is disagreement or which are not included in this policy:

1. When a parent wants an accommodation that is listed in the "Procedures for the Implementation of Accommodations for Students with Special Needs" and which the school is not willing to provide, the request must be made, in writing, to an appropriately qualified supervisory officer (or supervisory officers) designated by the director of education to make decisions on such requests. In the interest of consistency, the number of individuals making these decisions should be limited.

   For students in Independent Schools the request must be made to the Director of the Field Services Branch, Ministry of Education.

   For students in Native Schools the request must be made to the Director of the Curriculum and Assessment Policy Branch, Ministry of Education.

   For students in Provincial Schools the request must be made to the Director of the Provincial Schools Branch of the Ministry of Education.

2. When a parent and/or principal wants an accommodation that is not listed in the "Procedures for the Implementation of Accommodations for Students with Special Needs," and that conforms to the definition of accommodation in section II of this document, the request must be made, in writing, to the Chief Assessment Officer of the Education Quality and Accountability Office (EQAO).

Human Resources required for the implementation of accommodations

See Part VI, "Human Resources Required for the Implementation of Accommodations for Students with Special Needs," for direction on the qualifications and roles of individuals supporting the implementation of accommodations for students with special needs.

V Accommodations for Students with Special Needs

Note: To ensure the integrity of the test and the equitable participation of all students, only the accommodations listed below may be provided for students with special needs, unless procedures outlined in Part IV of this document ("Procedures for resolution of disagreements") apply. In each case, the accommodation used must be specified in the student's IEP.

Accommodations for the Reading component of the test

1. Setting of the test
   - Students may be provided an individual or small-group setting in a separate location or familiar space.
   - Individual study carrels may be provided.
   - Students may be given preferential seating within the regular classroom.
   - Special lighting may be provided.
   - Assistive devices or adaptive equipment (such as a device to support the student's arm for printing, writing or keyboarding, special pens or pencil grips) may be provided.
2. Timing
   - Additional time for the test beyond the time assigned is permitted, but not more than double the time. Each component must be completed by the end of each test day. Under the provisions of Part IV, "Procedures for resolution of disagreements" (item 2), a request

for consideration of more than double the time may be made to the Chief Assessment Officer of EQAO.

- Periodic supervised breaks may be provided. The duration of the break will not be deducted from the time allowed for the component of the test on which the student is working.
- In cases of students with severe attention problems, prompts may be given solely for the purpose of drawing the student's attention back to the test. (See Part VI for direction on the use of prompters.)

3. Presentation format
   - When reading from the teacher's administration script, teachers may not change or explain any of the instructions that they read to the students. They may repeat instructions.
   - Teachers may not read aloud, change or explain the instructions, texts or questions in the student booklets. (See Part VII for notes on pre-test availability of instructions that will appear in student booklets.)
   - Tests may be presented in Braille or large-print versions.
   - The use of American Sign Language or Signed Language or an oral interpreter is permitted for instructions contained in the teacher's administration script.
   - Coloured overlays and coloured paper are permitted.
   - Assistive devices, such as a Kurzweil reader, those permitting optical-to-tactile transformations and "universally scripted visual components," or other devices indicated by the IEP, are permitted.
   - A teacher (or other adult as outlined in Part VI) may assist the student in separating the test booklet in order that text and questions can be referred to side by side.
   - An electronic format may be used. The test will be available in formats compatible with a variety of platforms.
   - An audiotaped version of the test may be used.

   It will be necessary for principals to indicate on the *Student Verification Report*, which will be sent to the schools well in advance of the administration of the test, the number of Braille, large-print, electronic and audiotaped versions of the test the students will require. This will allow EQAO to supply the required versions of the test and to monitor their use across the province.

4. Response format
   - Provisions may be made for students who require extra space for their responses to the test questions and writing prompts.
     - The use of the following is permitted:
       - a Brailler
       - a speech synthesizer or Bliss board
       - voice-to-text software, such as Dragon Naturally Speaking

5. Response format: Accommodations Requiring Specific Documentation

When

i) the student has a special need that affects written expression and this special need is such that the student cannot translate thoughts into written language, even though the student can express thoughts orally, or

ii) the student has a special need that affects written output and the time constraint of the on-demand task will inhibit the required product, or

iii) the student has a special need that affects small motor control and his/her handwriting is illegible and would be unintelligible for the marker to make an appropriate assessment, and the accommodation is included in the student's IEP, the student may be provided with

> verbatim scribing of the responses or tape recording for later transcription,

or

> a computer or word processor.

In such cases, a form to be supplied by EQAO must be sent to the Chief Assessment Officer at the Education Quality and Accountability Office before the administration of the Grade 10 Test of Reading and Writing Skills. The information provided on this form will allow EQAO to monitor the use of verbatim scribers and computers/word processors across the province.

Please note that for the October 2000 administration of the Grade 10 Test of Reading and Writing Skills the requirement that a form be submitted to EQAO will be waived. All accommodations must be documented on the *Student Information Form*.

See Part VI for direction on scribing.

Accommodations for the Writing Component of the test

Please note that for many of the accommodations listed below, EQAO support is available.

1. Setting of the test
   - Students may be provided an individual or small-group setting in a separate location or a familiar space.
   - Individual study carrels may be provided.
   - Students may be given preferential seating within the regular classroom.
   - Special lighting may be provided.
   - Assistive devices or adaptive equipment (such as a device to support the student's arm for printing, writing or keyboarding or special pens or pencil grips) may be provided.

2. Timing
   - Additional time to complete the test beyond the time assigned is permitted but not more than double the time. Each component must be completed by the end of the test day.
   - Periodic supervised breaks are permitted. The duration of the break will not be deducted from the time allowed for the component of the test on which the student is working.
   - In cases of students with severe attention problems, prompts may be given solely for the purpose of drawing the student's attention back to the test. (See Part VI for notes on the use of prompters.)

3. Presentation format
    - When reading from the teacher's administration script, teachers may not change or explain any of the instructions that they read to the students. They may repeat instructions.
    - Verbatim reading of the writing prompts and accompanying texts is permitted where necessary. (See Part VII for notes on pre-test availability of instructions that will appear in student booklets.)
    - Tests may be presented on audiotape or in Braille or large print. A device permitting optical-to-tactile transformations and "universally scripted visual components" may be used.
    - Use of American Sign Language, Signed Language or an oral interpreter is permitted for instructions and writing prompts and accompanying texts.
    - Coloured overlays and coloured paper are permitted.
    - Assistive devices, such as a Kurzweil reader, or other assistive devices as indicated by the IEP, may be used.
    - A teacher (or other adult as outlined in Part VI) may assist the student in separating the test booklet in order that text and question can be referred to side by side.
    - An electronic format may be used. The test will be available in formats compatible with a variety of platforms.
4. Response format
    - The use of the following is permitted:
        - a Brailler
        - a speech synthesizer or Bliss board
        - voice-to-text software, such as Dragon Naturally Speaking
5. Response format: Accommodations Requiring Specific Documentation

    When

    i) the student has a special need that affects written expression and this special need is such that the student cannot translate thoughts into written language, even though the student can express thoughts orally, or

    ii) the student has a special need that affects written output and the time constraint of the on-demand task will inhibit the required product, or

    iii) the student has a special need that affects small motor control and his/her handwriting is illegible and would be unintelligible for the marker to make an appropriate assessment, and the accommodation is included in the student's IEP, the student may be provided with

    > verbatim scribing of the responses or tape recording for later transcription,

    or

    > a computer or word processor.

In such cases, a form to be supplied by EQAO must be sent to the Chief Assessment Officer at the Education Quality and Accountability Office before the administration of the Grade 10 Test of Reading and Writing Skills. The information provided on this form will allow EQAO to monitor the use of verbatim scribers and computers/word processors across the province.

For the October 2000 administration of the Grade 10 Test of Reading and Writing Skills, the requirement that a form be submitted to EQAO will be waived. All accommodations must be documented on the *Student Information Form*.

See Part VI for direction on scribing.

It will be noted on the student's *Individual Student Report* that spelling was not evaluated for students whose writing component of the test was scribed.

Students who use a keyboard to complete the test will be permitted to use spell-checking. It will be noted on the student's *Individual Student Report* that spelling was not evaluated.

The lack of a score for spelling will not compromise the student's chances of satisfactorily completing the test.

VI Human Resources Required for the Implementation of Accommodations for Students with Special Needs

Any person engaged to support the accommodations of one or more students, if not a teacher himself or herself, must work under the direct supervision of a teacher. Only adult persons who are not relatives of students writing the test may provide such support. It is important that training be provided to these individuals and that the following guidelines be strictly adhered to.

- The role of those supporting the implementation of accommodations for students with special needs is to ensure that the student is properly supervised during the administration of the test and that the accommodations are appropriately implemented.
- Individuals supporting the implementation of accommodations for students with special needs shall not provide students with any assistance that would compromise the validity of the test; that is, they shall not provide assistance that helps students to understand the questions or to formulate their answers or to give prompts about pacing their work.
- As a general principle, it is best that the person providing the support for the implementation of accommodations be a person who is known to the student, but not a relative.
- All non-teaching personnel providing support for the implementation of accommodations will be required to sign a confidentiality agreement and will be required to adhere to the same security provisions and other guidelines for the administration of the Grade 10 Test of Reading and Writing Skills as the teachers and principals involved in its administration.
- Scribes:

The role of the scribe is to write, type or word process the student's dictated answers to questions. The scribe must not in any way assist the student in formulating answers to the questions or writing prompts. The scribe must

> not give factual help to the student and must not offer any suggestions;

> not give advice to the student related to the order or pacing of the questions;

> write, type or word process answers exactly as they are dictated by the student;

> may read back what has been recorded at the student's request;

- Transcription of Tape Recorded or Brailled Student Responses:

  EQAO will arrange for the transcription of tape recorded or Brailled student responses.

- Prompters:

  Prompters are allowed for students with severe attention difficulties where the provision of supervised breaks will not adequately accommodate the student's needs. The sole purpose for prompting is to draw the student's attention back to the test through a tap on the student's arm or shoulder or on the desk or through a verbal signal. The prompter may not draw the attention of the student to a specific part of the student booklet. The prompter's role is to observe the student's behaviour, not to observe the quality or pace of the student's work.

VII Familiarizing Students with Test Format and Instructions Prior to the Test

Materials that resemble the student test booklets have been posted on the EQAO Web site. These materials can be used to familiarize all students with the test and will give opportunities for teachers to provide clarification and explanation for all students well in advance of the administration of the test. These materials give a clear indication of the types of tasks on the test. The language is similar to instructions students will find in their booklets.

Teachers have been encouraged to use the materials to familiarize students with the type of reading selections, the format and wording that they will find on the test. This way students can be well prepared for the test without compromising the integrity of the test itself.

Support Materials are available for all students participating in the Grade 10 Test of Reading and Writing Skills:

- In February 2000, a sample reading passage and sample questions (with expected responses and explanations), a list of the kinds of writing tasks (a summary, a paragraph, a short report and a series of paragraphs supporting an opinion) and a sample writing task were included in the Field Test Bulletin that was posted on the EQAO Web site.
- In April 2000, three "Fast Facts" information flyers (for Students, Parents/Guardians and Educators) were posted on the Web site to provide more information on what to expect on the test. Each "Fast Facts" sheet included a section on how to help students prepare for the test.
- In May 2000, a package of *Support Materials* for the reading component of the test was posted on the Web site. It offered "Tips for Students," including general advice on how to approach the reading questions and a glossary of terms frequently used in the questions. In addition, for each type of reading selection (informational text, graphical text and literary text), there were specific steps and strategies suggested for reading the texts and answering the questions.

- In early June 2000, a set of sample writing tasks was made available on the Web site. This posting included "Writing Tips for Students" which gave general suggestions related to the writing components of the test and more specific advice related to each of the four types of writing required on the test (a summary, a series of paragraphs supporting an opinion, a news report and an informational paragraph).
- The *Support Materials* were designed to reflect the actual test materials in their use of language, level of difficulty, type of text, type of questions or writing prompts, layout, overall format, etc. In short, working through the six selections and questions of the reading component of the package and the four writing tasks gives students a clear sense of what to expect during the actual testing experience.

Other supports available for all students participating in the test include the reading of a specially prepared teacher script which gives instructions to all students to supplement the written instructions contained in their booklets. The instructions contained in the teacher script may be repeated verbatim by the teacher upon request.

Appendix E

Special Provisions for Students Enrolled in ESL/ELD Programs

Students who have not yet acquired the level of proficiency in English required to successfully complete the test and who are enrolled in an English as a Second Language (ESL) or English Literacy Development (ELD) programs at levels 1, 2, 3 and 4 should be considered for a "deferral" of the test.

Ministry of Education policy[1] states that the principal, in consultation with the parent, student (or adult student) and appropriate school staff will determine whether or not a deferral should be granted. A deferral may be requested by a parent or adult student or the consideration of a deferral may be initiated by the principal. In cases where the parent or adult student disagrees with the principal, the parent or adult student may ask the appropriate supervisory officer to review the matter.

It may be advisable in some cases for students to attempt the test before they have completed ESL levels 1–4 or ELD levels 1–4. Special provisions may be put in place for students who are enrolled in an English as a Second Language (ESL) or English Literacy Development (ELD) program at levels 1, 2, 3 and 4 who write the Grade 10 Test of Reading and Writing Skills. Special provisions may also be put in place for students enrolled in a Bridge to English (ESL level 5) program.

Students who have completed ESL/ELD levels 1–4 and the Bridge to English courses will not be permitted special provisions.

In the Grade 10 Test of Reading and Writing Skills, special provisions must follow the EQAO guidelines as outlined below in order that the integrity and validity of the test is maintained. Strict adherence to the guidelines for special provisions and careful documentation of practices will ensure consistency and fairness across the province and from year to year.

Decisions about Special Provisions:

- Must always be made on an individual basis;
- Must be documented in full on the *Student Information Form*;
- Must be made in consultation with the student, parent, appropriate teaching staff and the principal; and
- Must be made prior to the administration of the test.

1 Ministry of Education and Training. (1999). *Ontario Secondary Schools Grades 9 to 12: Program and Diploma Requirements.* Queen's Printer for Ontario.

Special Provisions for ESL/ESD Students:
The following special provisions are permitted for **both the reading and writing components** of the test:

1. Timing:
   > Additional time for the test beyond the time assigned is permitted, but not more than double the time. Each component must be completed by the end of each test day.
   > Periodic supervised breaks may be provided. The duration of the break will not be deducted from the time allowed for the component of the test on which the student is working.
2. Setting:
   > Students may be provided an individual or small-group setting in a separate location or familiar space.
   > Individual study carrels may be provided.

Please note that guidelines for students who are enrolled in ESL/ELD programs and who also have special needs are outlined in the "Procedures for the Implementation of Accommodations for Students with Special Needs." These procedures pertain to ESL/ELD students with special needs.

Appendix F

Recommended Resources for the Ontario Secondary School Literacy Test[1]

Alderson, J.C., Clapham, C. and Wall, D. (1995). *Language test construction and evaluation.* Cambridge University Press: Cambridge.

Bachman, L.F. and Palmer, A.S. (1996). *Language testing in practice.* Oxford University Press: Oxford.

Bizzel, Patricia. (1982). Cognition, convention and certainty: What we need to know about writing. *Pre / Text,* 3.

Bereiter, C. and Scardamalia, M. (1987). An attainable version of high literacy: Approaches to teaching higher order skills in reading and writing. *Curriculum Inquiry,* 17:1.

Camp, Roberta. (1993). The place of portfolios in our changing views of writing assessment. *Construction versus Choice in Cognitive Measurement.* (ed. Randy Elliot Bennett, William C. Ward). Hillsdale, NJ: Lawrence Erlbaum Associates, 29–44, 183–212.

Emig, Janet. (1983). Literacy and freedom. *The web of meaning: Essays on writing, teaching, learning and thinking.* Upper Montclair, NJ: Boynton/Cook.

Gambell, Trevor & Hunter, Darryl M. (1999). Rethinking gender differences in literacy. *Canadian Journal of Education, 24,* 1, Winter.

Gipps, Caroline. (1994). *Beyond Testing.* Washington, DC: Falmer Press.

Giasson-Lachance, J. (1995). *La lecture: de la théorie à la pratique.* Montréal: G. Morin, ed.

Gitomer, Drew H. (1993). Performance assessment and educational measurement. *Construction versus choice in cognitive measurement.* Hillsdale, NJ and London: Lawrence Erlbaum Associates.

Haladyna, Thomas M. (1997). *Writing test items to evaluate higher order thinking.* Boston: Allyn and Bacon.

Hamp-Lyons, Liz. (1995). Uncovering possibilities for a constructivist paradigm for writing assessment. *College composition and communication, 46* (3).

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms.* Cambridge: Cambridge University Press.

Hunter, Darryl M., Jones, Richard M., & Bikkar, S. Randhawa. (1996). The use of holistic versus analytic scoring for large-scale assessment of writing. *Canadian Journal of Program Evaluation, 11,* 2, 61–85.

Jennings, Martha, Janna Fox, Barbara Graves & Elana Shohamy. (1999). The test-takers' choice: An investigation of the effect of topic on language test performance. *Language testing, 16,* 4, October, 426–456.

Laveault, Dany. (1999). The adaptation and use of tests in multiple languages and cultures. University of Ottawa: Faculty of Education. Commissioned by EQAO, Ontario.

Millman, Jason & Jennifer Greene. (1993). The specification and development of tests of achievement and ability. *Educational measurement* (3rd ed.) R. L. Linn (Ed.), New York: Macmillan.

Mosenthal, Peter B. (1998). Defining prose task characteristics for use in computer-adaptive testing and instruction. *American Educational Research Journal, 35,* 2, Summer 1998, 269–307.

Resnick, L.B. (1992). Assessing the thinking curriculum: New tools for educational reform. In Gifford B. (Ed.) *Changing assessments: Alternative views of aptitude, achievement and instruction.* London: Kluwer Academic Publishers.

Rosen, H. (1984). *Stories and meanings.* Sheffield, England: NATE Publications.

Speicher, Barbara L. & Jessica Bielanski. (2000). Critical thoughts on teaching standard English. *Curriculum Inquiry 30:2* (2000), Malden, MA and Oxford, UK: Blackwell Publishers.

Tardif, J. (1992). *Pour un enseignement stratégique: L'apport de la psychologie cognitive.* Montréal: Éditions Logiques.

Torrance, Nancy & David R. Olson. (1987). Development of the metalanguage and the acquisition of literacy: A progress report. *Interchange, 18,* 1/2 (Spring/Summer 1987), 136–146.

Traub, Ross E. (1993). On the equivalence of the traits assessed by multiple-choice and constructed-response tests. *Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment.* Hillsdale, NJ, and London: Lawrence Erlbaum Associates, Publishers.

Woolcot, Willa & Legg, Sue M. (1998). *An overview of writing assessment: Theory, research and practice.* Urbana, IL: National Council of Teachers of English.

Wiggins, Grant P. (1993). *Assessing student performance.* San Francisco: Jossey-Bass Publishers.

Zak, Frances & Weaver, Christopher. (1998). *The theory and practice of grading writing: Problems and possibilities.* Albany: State University of New York Press.

1 Please note that this resource list is not exhaustive. A more complete list will be posted on our Web site (www.eqao.com) at a later date.